

## RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number:

10/507,421

Source:

PGT

Date Processed by STIC:

3-28-05

# ***ENTERED***



PCT

## RAW SEQUENCE LISTING

DATE: 03/28/2005

PATENT APPLICATION: US/10/507,421

TIME: 14:23:50

Input Set : A:\2005-03-09 0760-0337PUS1.ST25.txt

Output Set: N:\CRF4\03282005\J507421.raw

4 &lt;110&gt; APPLICANT: NARIMATSU, Hisashi et al.

6 &lt;120&gt; TITLE OF INVENTION: NOVEL N-ACETYLGLUCOSAMINE TRANSFERASE, NUCLEIC ACID ENCODING

THE

7 SAME AND USE THEREOF IN DIAGNOSING CANCER AND/OR TUMOR

9 &lt;130&gt; FILE REFERENCE: 0760-0337PUS1

11 &lt;140&gt; CURRENT APPLICATION NUMBER: US 10/507,421

12 &lt;141&gt; CURRENT FILING DATE: 2004-09-13

14 &lt;150&gt; PRIOR APPLICATION NUMBER: PCT/JP03/03044

15 &lt;151&gt; PRIOR FILING DATE: 2003-03-14

17 &lt;150&gt; PRIOR APPLICATION NUMBER: JP 2002-70996

18 &lt;151&gt; PRIOR FILING DATE: 2002-03-14

20 &lt;160&gt; NUMBER OF SEQ ID NOS: 28

22 &lt;170&gt; SOFTWARE: PatentIn 3.2

24 &lt;210&gt; SEQ ID NO: 1

25 &lt;211&gt; LENGTH: 283

26 &lt;212&gt; TYPE: PRT

27 &lt;213&gt; ORGANISM: Homo sapiens

29 &lt;400&gt; SEQUENCE: 1

30 Tyr Phe Pro Met Leu Leu Asn His Pro Glu Lys Cys Arg Gly Asp Val

31 1 5 10 15

32 Tyr Leu Leu Val Val Lys Ser Val Ile Thr Gln His Asp Arg Arg

33 20 25 30

34 Glu Ala Ile Arg Gln Thr Trp Gly Arg Glu Arg Gln Ser Ala Gly Gly

35 35 40 45

36 Gly Arg Gly Ala Val Arg Thr Leu Phe Leu Leu Gly Thr Ala Ser Lys

37 50 55 60

38 Gln Glu Glu Arg Thr His Tyr Gln Gln Leu Leu Ala Tyr Glu Asp Arg

39 65 70 75 80

40 Leu Tyr Gly Asp Ile Leu Gln Trp Gly Phe Leu Asp Thr Phe Phe Asn

41 85 90 95

42 Leu Thr Leu Lys Glu Ile His Phe Leu Lys Trp Leu Asp Ile Tyr Cys

43 100 105 110

44 Pro His Val Pro Phe Ile Phe Lys Gly Asp Asp Asp Val Phe Val Asn

45 115 120 125

46 Pro Thr Asn Leu Leu Glu Phe Leu Ala Asp Arg Gln Pro Gln Glu Asn

47 130 135 140

48 Leu Phe Val Gly Asp Val Leu Gln His Ala Arg Pro Ile Arg Arg Lys

49 145 150 155 160

50 Asp Asn Lys Tyr Tyr Ile Pro Gly Ala Leu Tyr Gly Lys Ala Ser Tyr

51 165 170 175

52 Pro Pro Tyr Ala Gly Gly Gly Gly Phe Leu Met Ala Gly Ser Leu Ala

53 180 185 190

54 Arg Arg Leu His His Ala Cys Asp Thr Leu Glu Leu Tyr Pro Ile Asp

55 195 200 205

## RAW SEQUENCE LISTING

DATE: 03/28/2005

PATENT APPLICATION: US/10/507,421

TIME: 14:23:50

Input Set : A:\2005-03-09 0760-0337PUS1.ST25.txt

Output Set: N:\CRF4\03282005\J507421.raw

```

56 Asp Val Phe Leu Gly Met Cys Leu Glu Val Leu Gly Val Gln Pro Thr
57      210                      215                      220
58 Ala His Glu Gly Phe Lys Thr Phe Gly Ile Ser Arg Asn Arg Asn Ser
59 225                      230                      235                      240
60 Arg Met Asn Lys Glu Pro Cys Phe Phe Arg Ala Met Leu Val Val His
61                      245                      250                      255
62 Lys Leu Leu Pro Pro Glu Leu Leu Ala Met Trp Gly Leu Val His Ser
63                      260                      265                      270
64 Asn Leu Thr Cys Ser Arg Lys Leu Gln Val Leu
65      275                      280
67 <210> SEQ ID NO: 2
68 <211> LENGTH: 849
69 <212> TYPE: DNA
70 <213> ORGANISM: Homo sapiens
72 <400> SEQUENCE: 2
73 tac ttc ccc atg ctg ctg aac cac ccg gag aag tgc agg ggc gat gtc      48
74 Tyr Phe Pro Met Leu Leu Asn His Pro Glu Lys Cys Arg Gly Asp Val
75 1      5      10      15
76 tac ctg ctg gtg gtt gtc aag tcg gtc atc acg cag cac gac cgc cgc      96
77 Tyr Leu Leu Val Val Val Lys Ser Val Ile Thr Gln His Asp Arg Arg
78      20      25      30
79 gag gcc atc cgc cag acc tgg ggc cgc gag cgg cag tcc gcg ggt ggg      144
80 Glu Ala Ile Arg Gln Thr Trp Gly Arg Glu Arg Gln Ser Ala Gly Gly
81      35      40      45
82 ggc cga ggc gcc gtg cgc acc ctc ttc ctg ctg ggc acg gcc tcc aag      192
83 Gly Arg Gly Ala Val Arg Thr Leu Phe Leu Leu Gly Thr Ala Ser Lys
84      50      55      60
85 cag gag gag cgc acg cac tac cag cag ctg ctg gcc tac gaa gac cgc      240
86 Gln Glu Glu Arg Thr His Tyr Gln Gln Leu Leu Ala Tyr Glu Asp Arg
87 65      70      75      80
88 ctc tac ggc gac atc ctg cag tgg ggc ttt ctc gac acc ttc ttc aac      288
89 Leu Tyr Gly Asp Ile Leu Gln Trp Gly Phe Leu Asp Thr Phe Phe Asn
90      85      90      95
91 ctg acc ctc aag gag atc cac ttc ctc aag tgg ctg gac atc tac tgc      336
92 Leu Thr Leu Lys Glu Ile His Phe Leu Lys Trp Leu Asp Ile Tyr Cys
93      100      105      110
94 ccc cac gtc ccc ttc att ttc aaa ggc gac gat gac gtc ttc gtc aac      384
95 Pro His Val Pro Phe Ile Phe Lys Gly Asp Asp Asp Val Phe Val Asn
96      115      120      125
97 ccc acc aac ctg cta gaa ttt ctg gct gac cgg cag cca cag gaa aac      432
98 Pro Thr Asn Leu Leu Glu Phe Leu Ala Asp Arg Gln Pro Gln Glu Asn
99      130      135      140
100 ctg ttc gtg ggc gat gtc ctg cag cac gct cgg ccc att cgc agg aaa      480
101 Leu Phe Val Gly Asp Val Leu Gln His Ala Arg Pro Ile Arg Arg Lys
102 145      150      155      160
103 gac aac aaa tac tac atc ccg ggg gcc ctg tac ggc aag gcc agc tat      528
104 Asp Asn Lys Tyr Tyr Ile Pro Gly Ala Leu Tyr Gly Lys Ala Ser Tyr
105      165      170      175
106 ccg ccg tat gca ggc ggc ggt ggc ttc ctc atg gcc ggc agc ctg gcc      576

```

## RAW SEQUENCE LISTING

DATE: 03/28/2005

PATENT APPLICATION: US/10/507,421

TIME: 14:23:50

Input Set : A:\2005-03-09 0760-0337PUS1.ST25.txt

Output Set: N:\CRF4\03282005\J507421.raw

```

107 Pro Pro Tyr Ala Gly Gly Gly Gly Phe Leu Met Ala Gly Ser Leu Ala
108      180      185      190
109 cgg cgc ctg cac cat gcc tgc gac acc ctg gag ctc tac ccg atc gac      624
110 Arg Arg Leu His His Ala Cys Asp Thr Leu Glu Leu Tyr Pro Ile Asp
111      195      200      205
112 gac gtc ttt ctg ggc atg tgc ctg gag gtg ctg ggc gtg cag ccc acg      672
113 Asp Val Phe Leu Gly Met Cys Leu Glu Val Leu Gly Val Gln Pro Thr
114      210      215      220
115 gcc cac gag ggc ttc aag act ttc ggc atc tcc cgg aac cgc aac agc      720
116 Ala His Glu Gly Phe Lys Thr Phe Gly Ile Ser Arg Asn Arg Asn Ser
117 225      230      235      240
118 cgc atg aac aag gag ccg tgc ttt ttc cgc gcc atg ctc gtg gtg cac      768
119 Arg Met Asn Lys Glu Pro Cys Phe Phe Arg Ala Met Leu Val Val His
120      245      250      255
121 aag ctg ctg ccc cct gag ctg ctc gcc atg tgg ggg ctg gtg cac agc      816
122 Lys Leu Leu Pro Pro Glu Leu Leu Ala Met Trp Gly Leu Val His Ser
123      260      265      270
124 aat ctc acc tgc tcc cgc aag ctc cag gtg ctc      849
125 Asn Leu Thr Cys Ser Arg Lys Leu Gln Val Leu
126      275      280
128 <210> SEQ ID NO: 3
129 <211> LENGTH: 327
130 <212> TYPE: PRT
131 <213> ORGANISM: Homo sapiens
133 <400> SEQUENCE: 3
134 Ala Ser Gln Gly Pro Gln Ala Trp Asp Val Thr Thr Thr Asn Cys Ser
135 1      5      10      15
136 Ala Asn Ile Asn Leu Thr His Gln Pro Trp Phe Gln Val Leu Glu Pro
137      20      25      30
138 Gln Phe Arg Gln Phe Leu Phe Tyr Arg His Cys Arg Tyr Phe Pro Met
139      35      40      45
140 Leu Leu Asn His Pro Glu Lys Cys Arg Gly Asp Val Tyr Leu Leu Val
141      50      55      60
142 Val Val Lys Ser Val Ile Thr Gln His Asp Arg Arg Glu Ala Ile Arg
143 65      70      75      80
144 Gln Thr Trp Gly Arg Glu Arg Gln Ser Ala Gly Gly Gly Arg Gly Ala
145      85      90      95
146 Val Arg Thr Leu Phe Leu Leu Gly Thr Ala Ser Lys Gln Glu Glu Arg
147      100      105      110
148 Thr His Tyr Gln Gln Leu Leu Ala Tyr Glu Asp Arg Leu Tyr Gly Asp
149      115      120      125
150 Ile Leu Gln Trp Gly Phe Leu Asp Thr Phe Phe Asn Leu Thr Leu Lys
151      130      135      140
152 Glu Ile His Phe Leu Lys Trp Leu Asp Ile Tyr Cys Pro His Val Pro
153 145      150      155      160
154 Phe Ile Phe Lys Gly Asp Asp Val Phe Val Asn Pro Thr Asn Leu
155      165      170      175
156 Leu Glu Phe Leu Ala Asp Arg Gln Pro Gln Glu Asn Leu Phe Val Gly
157      180      185      190

```

## RAW SEQUENCE LISTING

DATE: 03/28/2005

PATENT APPLICATION: US/10/507,421

TIME: 14:23:50

Input Set : A:\2005-03-09 0760-0337PUS1.ST25.txt

Output Set: N:\CRF4\03282005\J507421.raw

```

158 Asp Val Leu Gln His Ala Arg Pro Ile Arg Arg Lys Asp Asn Lys Tyr
159      195      200      205
160 Tyr Ile Pro Gly Ala Leu Tyr Gly Lys Ala Ser Tyr Pro Pro Tyr Ala
161      210      215      220
162 Gly Gly Gly Gly Phe Leu Met Ala Gly Ser Leu Ala Arg Arg Leu His
163 225      230      235      240
164 His Ala Cys Asp Thr Leu Glu Leu Tyr Pro Ile Asp Asp Val Phe Leu
165      245      250      255
166 Gly Met Cys Leu Glu Val Leu Gly Val Gln Pro Thr Ala His Glu Gly
167      260      265      270
168 Phe Lys Thr Phe Gly Ile Ser Arg Asn Arg Asn Ser Arg Met Asn Lys
169      275      280      285
170 Glu Pro Cys Phe Phe Arg Ala Met Leu Val Val His Lys Leu Leu Pro
171      290      295      300
172 Pro Glu Leu Leu Ala Met Trp Gly Leu Val His Ser Asn Leu Thr Cys
173 305      310      315      320
174 Ser Arg Lys Leu Gln Val Leu
175      325
177 <210> SEQ ID NO: 4
178 <211> LENGTH: 981
179 <212> TYPE: DNA
180 <213> ORGANISM: Homo sapiens
182 <400> SEQUENCE: 4
183 gcc tct cag ggg ccc cag gcc tgg gac gtg acc acc act aac tgc tca      48
184 Ala Ser Gln Gly Pro Gln Ala Trp Asp Val Thr Thr Thr Asn Cys Ser
185 1      5      10      15
186 gcc aat atc aac ttg acc cac cag ccc tgg ttc cag gtc ctg gag ccg      96
187 Ala Asn Ile Asn Leu Thr His Gln Pro Trp Phe Gln Val Leu Glu Pro
188      20      25      30
189 cag ttc cgg cag ttt ctc ttc tac cgc cac tgc cgc tac ttc ccc atg      144
190 Gln Phe Arg Gln Phe Leu Phe Tyr Arg His Cys Arg Tyr Phe Pro Met
191      35      40      45
192 ctg ctg aac cac ccg gag aag tgc agg ggc gat gtc tac ctg ctg gtg      192
193 Leu Leu Asn His Pro Glu Lys Cys Arg Gly Asp Val Tyr Leu Leu Val
194      50      55      60
195 gtt gtc aag tcg gtc atc acg cag cac gac cgc cgc gag gcc atc cgc      240
196 Val Val Lys Ser Val Ile Thr Gln His Asp Arg Arg Glu Ala Ile Arg
197 65      70      75      80
198 cag acc tgg ggc cgc gag cgg cag tcc gcg ggt ggg ggc cga ggc gcc      288
199 Gln Thr Trp Gly Arg Glu Arg Gln Ser Ala Gly Gly Gly Arg Gly Ala
200      85      90      95
201 gtg cgc acc ctc ttc ctg ctg ggc acg gcc tcc aag cag gag gag cgc      336
202 Val Arg Thr Leu Phe Leu Leu Gly Thr Ala Ser Lys Gln Glu Glu Arg
203      100      105      110
204 acg cac tac cag cag ctg ctg gcc tac gaa gac cgc ctc tac ggc gac      384
205 Thr His Tyr Gln Gln Leu Leu Ala Tyr Glu Asp Arg Leu Tyr Gly Asp
206      115      120      125
207 atc ctg cag tgg ggc ttt ctc gac acc ttc ttc aac ctg acc ctc aag      432
208 Ile Leu Gln Trp Gly Phe Leu Asp Thr Phe Phe Asn Leu Thr Leu Lys

```

## RAW SEQUENCE LISTING

DATE: 03/28/2005

PATENT APPLICATION: US/10/507,421

TIME: 14:23:50

Input Set : A:\2005-03-09 0760-0337PUS1.ST25.txt

Output Set: N:\CRF4\03282005\J507421.raw

```

209      130      135      140
210 gag atc cac ttc ctc aag tgg ctg gac atc tac tgc ccc cac gtc ccc      480
211 Glu Ile His Phe Leu Lys Trp Leu Asp Ile Tyr Cys Pro His Val Pro
212 145      150      155      160
213 ttc att ttc aaa ggc gac gat gac gtc ttc gtc aac ccc acc aac ctg      528
214 Phe Ile Phe Lys Gly Asp Asp Asp Val Phe Val Asn Pro Thr Asn Leu
215      165      170      175
216 cta gaa ttt ctg gct gac cgg cag cca cag gaa aac ctg ttc gtg ggc      576
217 Leu Glu Phe Leu Ala Asp Arg Gln Pro Gln Glu Asn Leu Phe Val Gly
218      180      185      190
219 gat gtc ctg cag cac gct cgg ccc att cgc agg aaa gac aac aaa tac      624
220 Asp Val Leu Gln His Ala Arg Pro Ile Arg Arg Lys Asp Asn Lys Tyr
221      195      200      205
222 tac atc ccg ggg gcc ctg tac ggc aag gcc agc tat ccg ccg tat gca      672
223 Tyr Ile Pro Gly Ala Leu Tyr Gly Lys Ala Ser Tyr Pro Pro Tyr Ala
224      210      215      220
225 ggc ggc ggt ggc ttc ctc atg gcc ggc agc ctg gcc cgg cgc ctg cac      720
226 Gly Gly Gly Gly Phe Leu Met Ala Gly Ser Leu Ala Arg Arg Leu His
227 225      230      235      240
228 cat gcc tgc gac acc ctg gag ctc tac ccg atc gac gac gtc ttt ctg      768
229 His Ala Cys Asp Thr Leu Glu Leu Tyr Pro Ile Asp Asp Val Phe Leu
230      245      250      255
231 ggc atg tgc ctg gag gtg ctg ggc gtg cag ccc acg gcc cac gag ggc      816
232 Gly Met Cys Leu Glu Val Leu Gly Val Gln Pro Thr Ala His Glu Gly
233      260      265      270
234 ttc aag act ttc ggc atc tcc cgg aac cgc aac agc cgc atg aac aag      864
235 Phe Lys Thr Phe Gly Ile Ser Arg Asn Arg Asn Ser Arg Met Asn Lys
236      275      280      285
237 gag ccg tgc ttt ttc cgc gcc atg ctc gtg gtg cac aag ctg ctg ccc      912
238 Glu Pro Cys Phe Phe Arg Ala Met Leu Val Val His Lys Leu Leu Pro
239      290      295      300
240 cct gag ctg ctc gcc atg tgg ggg ctg gtg cac agc aat ctc acc tgc      960
241 Pro Glu Leu Leu Ala Met Trp Gly Leu Val His Ser Asn Leu Thr Cys
242 305      310      315      320
243 tcc cgc aag ctc cag gtg ctc      981
244 Ser Arg Lys Leu Gln Val Leu
245      325
247 <210> SEQ ID NO: 5
248 <211> LENGTH: 1206
249 <212> TYPE: DNA
250 <213> ORGANISM: Homo sapiens
252 <400> SEQUENCE: 5
253 atg tcg ctg tgg aag aaa acc gtc tac cgg agt ctg tgc ctg gcc ctg      48
254 Met.Ser Leu Trp Lys Lys Thr Val Tyr Arg Ser Leu Cys Leu Ala Leu
255      1      5      10      15
256 gcc ctg ctc gtg gcc gtg acg gtg ttc caa cgc agt ctc acc cct ggt      96
257 Ala Leu Leu Val Ala Val Thr Val Phe Gln Arg Ser Leu Thr Pro Gly
258      20      25      30
259 cag ttt ctg cag gag cct ccg cca ccc acc ctg gag cca cag aag gcc      144

```

VERIFICATION SUMMARY

DATE: 03/28/2005

PATENT APPLICATION: US/10/507,421

TIME: 14:23:51

Input Set : A:\2005-03-09 0760-0337PUS1.ST25.txt

Output Set: N:\CRF4\03282005\J507421.raw